



CDLR Publications

Aspect: Access to Scottish Parliamentary Election Candidate Materials

Alan Dawson

Centre for Digital Library Research, University of Strathclyde, Glasgow G1 1XH

December 2002

alan.dawson@strath.ac.uk

This article was published in [Sconul Newsletter](#) 27, 2002

Background to the collection

At 11:52 on the morning of Wednesday 12 May 1999, Dr Winnie Ewing stood in front of her new colleagues and spoke the following words:

...the Scottish Parliament, which adjourned on 25 March 1707, is hereby reconvened.

As the oldest qualified member of the new Scottish Parliament, Dr Ewing was granted the honour of making the opening speech¹. The first Scottish parliamentary proceedings of the industrial age had begun at 9:30 the same day, but these historic words could not be spoken until all 129 members had been sworn in and the Parliament really had been convened. This moment was the culmination of a process that can be traced back to the establishment of the Scottish Home Rule Association in 1886, but it was the election in 1997 of a Labour government committed to devolution that finally paved the way for a referendum and the elections of 6 May 1999.

Capturing history

In recognition of the significance of the first elections to a Scottish parliament for almost 300 years, the Andersonian Library at the University of Strathclyde started a campaign to collect literature from the election - the newsletters, leaflets, postcards and other ephemera produced by candidates and political parties. Staff and students of the University were encouraged to intercept these materials on their journey from letterbox to bin, and to bring them to the Library for archiving and future research. This initiative was never going to be 100% successful, but the Library did manage to acquire literature from around 70% of constituency candidates (234 out of 339) and for 18 of the 19 registered parties that contested the election.

There are problems with holding and displaying this kind of material. Items are:

- produced in all sorts of shapes and sizes
- flimsy and easily damaged
- small and easily slipped into a bag or folder
- too ephemeral to catalogue individually, yet too heterogeneous to catalogue collectively

The collection was therefore an obvious candidate for digitisation, and that is what the Aspect project set out to do. All parties were contacted for permission to digitise their election literature, and none refused, with most being enthusiastic about the process.

As well as helping to conserve the original materials, the major advantage of digitisation is in making the collection readily accessible to a much wider audience. Although available to everyone, it is of particular relevance to students and researchers of political science, political history, Scottish affairs and communication, as well as to politicians, civil servants, journalists and librarians.

The digitisation process

Conceptually and technically, digitisation is a quick and simple process. Standard desktop scanners are adequate for many applications, while larger scanners and high-quality digital cameras are available in media services units. In practice, things are rarely as simple as they at first appear. The main problem with digitising the Aspect collection was the size of the task. The 234 candidates (from 73 constituencies) produced an average of two double-sided leaflets each, which required almost 1000 separate scans and image files. Then there was the material produced for the eight electoral regions, where the introduction of proportional representation attracted plenty of independent candidates and single-issue parties. And then there was the general literature produced by parties specifically for the election but not tied to any candidate or region.

To complicate matters further, some parties contested only regions but not constituencies, and vice versa, some individual candidates stood in both constituencies and regions, some were elected from regions without standing in constituencies, and some party literature combined coverage of the parliamentary election with candidates for the local government elections, which were taking place on the same day. Confused electors were faced with casting three votes (local, constituency and region), and confused digitisers were faced with identifying and capturing two of these three (constituency and region). The local elections might have been interesting but they were hardly historic, and you have to draw the line somewhere.

The critical issues for digitisation were therefore not the size, resolution or format of the images, but the selection, labelling and management of the digital collection. For the record, images were captured in colour in TIFF format at 150 pixels per inch, then scaled down and converted to JPEG format for screen display, reducing the size of an average A4 image from around 6.5Mb to 100Kb.

In order to ensure consistent and accurate identification of over 1200 images, a strict file-naming system was used, comprising a party code, constituency or region code, item number and page number. Thumbnail images were created for every item, immediately doubling the number of images to be handled.

Issues and promises

As well as creating digital images of the source material, the initial aim was to include transcriptions of all the text too. This turned out to be over-ambitious, as the process of optical character reading (OCR) required far more time than simple image capture, as all text had to be carefully checked for accuracy, and the colourful nature of the leaflets meant that many images had to be converted to black and white and further manipulated before OCR was at all feasible. In any case, transcribing every item would have been a waste of time, as there was substantial duplication of content. All the major parties had standard text which appeared time and again in leaflets for different candidates. OCR therefore needed to be selective, capturing only the original thoughts, policies and promises of individual candidates, and coverage of local issues for local people.

Even with this reduced scope it has not yet been feasible to transcribe all materials, but the content already available in searchable text form makes it possible to find out the relative importance given to different issues. For example, over 50% of constituency candidates made mention of the popular topics health or education, 18% covered employment, 15% mentioned drugs, 10% referred to the environment, 3% to racism and 1% to nuclear weapons. Analysing the coverage of education more specifically, 33% mentioned schools, 5% universities, 3% libraries and only 2% colleges, while the specific topic of tuition fees received 13% coverage. Only one candidate promised to draw the average wage and give the rest of his salary away, and only one candidate promised to buy a new cloak for the opening of the parliament.

Displaying the collection

One of the beauties of managing a digital collection, as opposed to a physical collection, is that once you have captured the content and described it accurately, it becomes possible to organise and access it in a number of ways, to suit users with different requirements. The substantive content of the Aspect collection is a separate web page for each constituency candidate, with a thumbnail for every available image. Users may reach this by starting from the Aspect home page² and browsing via one of four indexes:

- Constituency name
- Party name
- Candidate name
- Electoral region (if the candidate stood in both constituency and region)

As well as capturing the campaign literature, the full election results were collected and integrated with the source material, so it is possible to see at a glance how each candidate fared in a given constituency, or how a given party fared in each constituency. Even where no digital images are available, there is still a page displaying the basic election data for every constituency candidate. This completeness of coverage and level of integration has made the Aspect collection into much more than a set of digital objects, and offers far more to users than the physical literature collection.

Cataloguing the collection

In view of the structured nature of the materials, it was not necessary to carry out cataloguing as a separate exercise from content creation. By gathering the full set of candidate, party, constituency and region names and results from official sources³ and loading this into a database, all the information required for a catalogue record was available, allowing automated creation of records to be generated for different purposes:

- Metadata in Dublin Core format, embedded within each web page, to enable discovery by Internet search engines and harvesting by collaborative projects.
- A single file of records in Marc 21 format for uploading to a Horizon database, to facilitate cross-searching with other local and distributed collections.
- A set of related files in Marc 21 format for uploading to the OCLC Connexion service⁴, to assist international attempts at improving description and discovery of Internet resources.

This flexibility was achieved by entering details of the Marc codes into an Access database along with the structured data and transcribed text, from where it could be extracted in various ways by a purpose-written Visual Basic program. To complete the content management process, entering the filenames of digital images associated with each record into the same database enabled almost all of the Aspect website to be generated automatically.

This approach proved to be simple and robust enough to be adopted by a larger digitisation project⁵ as an interim measure, pending purchase and implementation of a special-purpose content management system.

Context

Although usable as a resource in its own right, Aspect is just one of a number of components of the Glasgow Digital Library (GDL)⁶. As well as creating and managing original content, a crucial part of the methodology of the GDL is to develop an information environment that will:

- ensure compatibility with international standards for resource description and cross-searching.
- enable GDL resources to be visible world-wide and discoverable by those with no prior knowledge of their existence.
- allow collections of digital objects to be assembled and presented in different ways.
- illustrate mechanisms for building large virtual collections from small distributed ones.

The process of creating and maintaining digital content has provided essential practical experience of applying standards to collection management and website development, thereby combining theory with practice and enabling a research project to become a user service. Having an infrastructure and methodology in place means that it should be possible to plug in new digital collections relatively easily and cheaply, while maintaining compliance with established standards.

Future developments

The 1999 election was unique in being the first Scottish parliamentary election for 300 years, but there is another one due in 2003 that could be recorded in similar fashion, should funding for digitisation be secured. The months leading up to the next election may be a good time to publicise the Aspect collection, as it provides a definitive record of 'what they said last time' for those seeking re-election. The Scottish Executive is keen to encourage greater interest in elections at all levels, and the broader issue of citizenship is one of a number of current priorities for the Scottish school curriculum.

There are rich possibilities for adding value to the Aspect collection by using it as the basis for structured learning materials and for involving more electors in the democratic process. The content may not have the high profile or immediate impact of a televised election broadcast, but it has a much wider potential audience and a much longer life span. Politicians don't always like to be reminded of their previous policies and promises. Now librarians, or indeed anyone with Internet access, will be able to

remind them. That is something the Scottish parliamentarians of 1707 never had to worry about.

References

1. A full transcript of the opening proceedings is available at:
http://www.scottish.parliament.uk/official_report/session99-00/or010104.htm
 2. Aspect is available via <http://gdl.cdrl.strath.ac.uk/aspect/>
 3. A detailed analysis of the election results is available in House of Commons research paper 99/50, available at <http://www.parliament.uk/commons/lib/research/rp99/rp99-050.pdf>. This provides statistics about parties and constituencies, but names of candidates were assembled from a range of other sources, including the source material for the Aspect collection.
 4. OCLC Connexion aims to integrate all cataloguing functions into a single interface that can be accessed from either a browser or a future Windows client. It is based on the OCLC CORC, CatExpress, CatME and Passport services, and is available via <http://connexion.oclc.org/>
 5. The Victorian Times project offers an examination of the social, political, and economic development of Great Britain in the Victorian era. Initial digitised content is available via <http://victoria.cdrl.strath.ac.uk/>
 6. The Glasgow Digital Library is a collaborative project that seeks to create, collect and provide access to digital content that is *about* Glasgow, *in* Glasgow or *for* Glasgow. The initial GDL project was funded for two years by the Research Support Libraries Programme. The GDL is now available as a user service via <http://gdl.cdrl.strath.ac.uk/>
-